

# 面向交通标志检测的空间信息聚合网络

张俊龙, 韩文花\*

(上海电力大学自动化工程学院, 上海 200090)

**摘要:** [目的] 针对日常交通驾驶过程中交通标志检测的漏检、误检以及模型体积较大等问题, 提出一种空间信息聚合网络 EDR-YOLOv8 算法。[方法] 通过增强边缘信息、多尺度上下文语义信息的捕获以及浅层与深层特征的再融合, 完成空间信息的聚合。为保留较多的交通标志边缘特征, 提出结合 Sobel 算子的边缘检测模块; 为捕捉更多的多尺度上下文语义信息, 设计膨胀残差修复块, 并将其集成到 C2f(faster implementation of CSP bottleneck with 2 convolutions) 中得到多尺度感知模块 DRRC; 为加强高、低层信息交互并重新校准目标位置, 推出重校准内容引导金字塔网络。通过对锚框与真实框不匹配方向的考虑和更关注普通质量的锚框, 设计 Wise-SIoU。[结果] 在 TT100K 数据集上,  $P$ 、交并比阈值 0.50 的平均精度均值、交并比阈值在 [0.50, 0.95] 区间的平均精度均值分别提升 12.5%、8.9%、10.3%, 且参数量下降了 23%。[结论] 本文方法相较原模型及其他算法, 表现出了精度较高和参数量较少的优点, 改善了漏检和误检问题, 并有利于边缘设备的部署。

**关键词:** 目标检测; 交通标志; YOLOv8; 深度学习; Wise-SIoU

中图分类号: TP 391.4; U 495

文献标志码: A

文章编号: 0438-0479(2025)04-0616-13

## Spatial information aggregation network for traffic sign detection

ZHANG Junlong, HAN Wenhua\*

(College of Automation Engineering, Shanghai University of Electric Power, Shanghai 200090, China)

**Abstract:** [Objective] In response to issues of missed detections, false detections, and large model size in traffic sign detection during daily driving, an improved EDR-YOLOv8 algorithm is proposed. [Methods] This algorithm enhances the capture of edge information and multi-scale contextual semantic information, while achieving spatial information aggregation through the re-fusion of low-level and high-level features. To retain more edge features of traffic signs, we propose an edge detection module (EDB) by combining the Sobel operator. Additionally, for the purpose of capturing sufficient multi-scale contextual semantic information, the dilated residual repair (DRR) block is designed and integrated into the C2f module, resulting in the multi-scale perception module DRRC. To enhance the interaction between high-level and low-level information and to recalibrate target locations, we also propose RCG-FPN. Furthermore, a Wise-SIoU is designed by considering the direction of mismatch between anchor boxes and ground truth boxes and placing greater emphasis on anchors of ordinary quality. [Results] On the TT100K dataset, the precision ( $P$ ), mAP@0.5, and mAP@0.5:0.95 increased by 12.5%, 8.9%, and 10.3% respectively, while the number of parameters decreased by 23%. [Conclusions] Compared to the original model and other algorithms, the proposed method exhibits advantages in higher accuracies and reduced parameter count, thus improving those issues of missed and false detections and facilitating the deployment on edge devices.

**Keywords:** object detection; traffic sign; YOLOv8; deep learning; Wise-SIoU

交通标志为驾驶员和自动驾驶车辆提供了如道路 规章、限速要求、方向指示等关键信息, 准确和高效的交

收稿日期: 2024-09-28 录用日期: 2025-04-28

\* 通信作者: hanwenhua@shiep.edu.cn

引文格式: 张俊龙, 韩文花. 面向交通标志检测的空间信息聚合网络[J]. 厦门大学学报(自然科学版), 2025, 64(4): 616-628.

Citation: ZHANG J L, HAN W H. Spatial information aggregation network for traffic sign detection[J]. J Xiamen Univ Nat Sci, 2025, 64(4): 616-628. (in Chinese)



通标志检测不但能够提高驾驶安全,还能增强自动驾驶系统的决策能力.深度学习在目标检测领域发展迅速,以 Faster R-CNN (faster region based-convolutional neural network)<sup>[1]</sup> 和 Mask R-CNN (mask region based-convolutional neural network)<sup>[2]</sup> 为代表的两阶段算法首先利用区域提议网络从输入图像中提取候选区域,再在这些假设区域上进行分类和边界框回归.这类方法精度较高,但训练过程复杂、速度较慢.而一阶段方法中的 YOLO (you only look once)<sup>[3-4]</sup> 和 SSD (single shot multiBox detector)<sup>[5]</sup> 直接在全图像上进行检测,将目标检测简化为一个回归问题,在保持较高的运行速度的同时,检测精度逐渐接近两阶段方法.

井方科等<sup>[6]</sup>通过对检测层改进并设计了双头检测结构,聚焦小目标特征信息的同时减少参数量.朱宁可等<sup>[7]</sup>对 Mobilenetv3 中倒残差结构进行改进,并取代 YOLOv5 的主干网络,使其更符合轻量网络的设计;高良鹏等<sup>[8]</sup>利用 C2f (faster implementation of CSP bottleneck with 2 convolutions) 和 FasterNet 的高效融合完成对 YOLOv8 网络的轻量化.上述研究网络模型在轻量化上有一定改进,但在检测精度上有所不足.

YOLOv8 是一个实时高效的目标检测模型,旨在支持物体分类、检测和分割等多种任务,并因其高效性和准确性而备受瞩目.为了解决一阶段算法在交通标志检测过程中检测精度不高、定位不准和复杂场景下难以识别的问题,基于 YOLOv8,提出一种具有高精度的轻量级的空间信息聚合网络 EDR-YOLOv8,具体如下:

1) 边缘检测模块(edge detection block, EDB):设计 Sobel 算子+最大池化的 EDB 替换原标准特征提取模块 CBS(convolution-batchnorm-silu),以实现边缘信息的高效提取;

2) 多尺度感知模块:设计具有多感受野的跨阶段膨胀残差修复块(dilation residual and repair module-C2f, DRRC)改进 C2f 模块,用于捕捉不同尺度的上下文信息;

3) 颈部网络:设计重校准内容引导的特征金字塔网络 (recalibrate content guided-feature pyramid network, RCG-FPN) 改进空间金字塔快速池化 (spatial pyramid pooling fast, SPPF) 模块,用于聚合边界信息和语义信息以及低级特征与高级特征的融合;

4) 损失函数:设计 Wise-SIoU, 增强对普通质

量样本的关注和对锚框与真实框不匹配方向的惩罚.

## 1 空间信息聚合网络 EDR-YOLOv8

针对交通标志检测过程中存在的精度不高、模型较大的问题,本文在 YOLOv8 算法的基础上,提出如图 1 所示的空间信息聚合网络 EDR-YOLOv8. 首先,为更好地保留交通标志边缘特征,EDB 使用 Sobel 算子提取水平和垂直向的突变信息,同时结合最大池化捕获空间信息,进行降维.其次,为解决交通标志多尺度特征提取问题,融合膨胀残差模块 (dilation-wise residual module, DWR)<sup>[9]</sup> 和膨胀重参块 (dilated reparam block, DRB)<sup>[10]</sup> 得到膨胀残差修复块 (dilation residual and repair module, DRR),旨在从图像中捕获多尺度上下文信息,并将其应用到 C2f (faster implementation of CSP bottleneck with 2 convolutions) 中得到 DRRC. 为了加强浅层和深层特征的信息交互能力,推出 RCG-FPN, 其使用选择性边界聚合 (selective boundary aggregation, SBA) 模块<sup>[11]</sup> 聚合边界信息和语义信息,再使用内容引导注意力融合 (content-guided attention fusion, CGAFusion) 模块<sup>[12]</sup> 充分引导模型关注每个通道的重要区域,保证低级特征与高级特征的充分融合.最后,设计 Wise-SIoU, 重点关注普通质量的锚框,同时考虑锚框与真实框不匹配的类型,以提高模型的定位性能. EDR-YOLOv8 网络结构如图 1 所示.

### 1.1 EDB

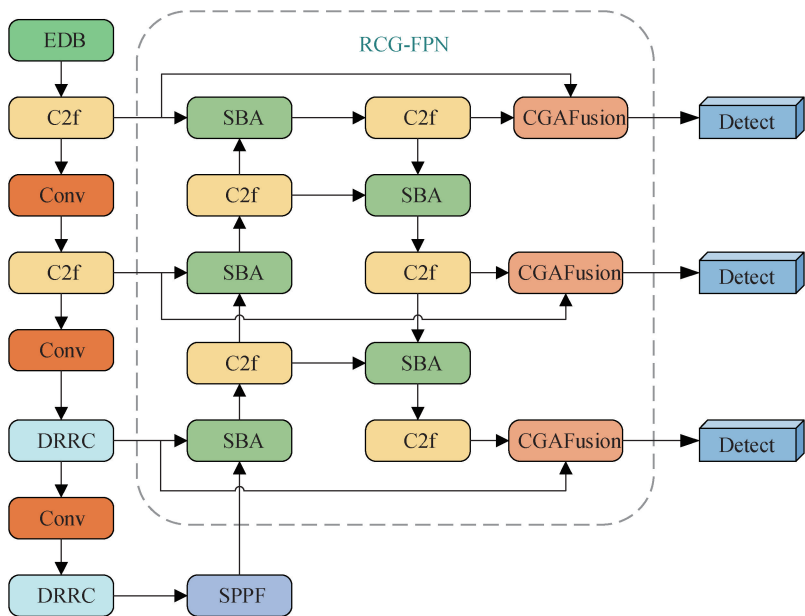
交通标志通常有清晰的形状轮廓和边缘,通过检测这些边缘有助于定位标志的位置,可以为后续的分类和识别提供准确的输入,并帮助区分目标和背景. EDB 结构如图 2 所示. EDB 运算过程为:

$$\mathbf{F}_1 = C_3(\mathbf{X}), \quad (1)$$

$$\mathbf{F}_2 = C_1(C_3(f_{MP}(\mathbf{F}_1) \oplus f_{sobel}(\mathbf{F}_1))), \quad (2)$$

式中,  $\mathbf{X}$  为输入,  $\mathbf{F}_1$  为经卷积层处理过的初级特征,  $\oplus$  为拼接操作,  $C_k$  为  $k \times k$  大小的卷积层,  $f_{MP}$  为最大池化操作,  $f_{sobel}$  为 Sobel 操作,  $\mathbf{F}_2$  为输出特征.

EDB 首先通过一个下采样卷积来处理交通标志图像提取初始特征,保留边缘、角落和纹理等低级特征的同时减小图像大小.再分为两个分支进行处理:一个使用 Sobel 算子检测水平和垂直方向的边缘,以检测图像中强度的突然变化和图像中交通标志字母、符号或边框的边缘的特征;另一分支使用最大池化捕获局部区域中最突出的特征,提取并保留重要的空间



SPPF(spatial pyramid pooling fast):空间金字塔池化.

图 1 EDR-YOLOv8 网络结构

Fig. 1 The network structure of EDR-YOLOv8

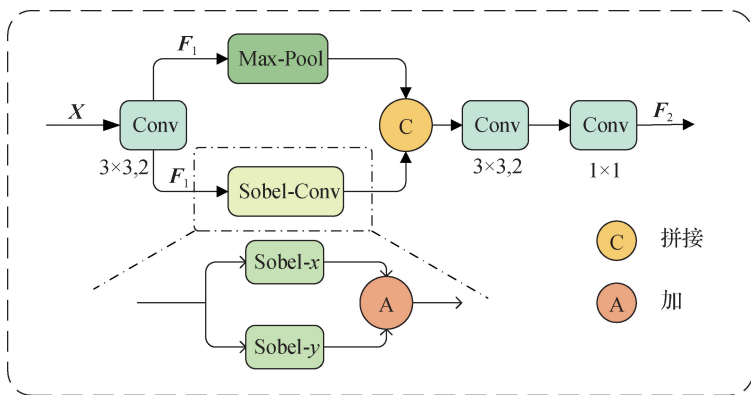


图 2 EDB 结构

Fig. 2 EDB structure

信息.接着,通过卷积层将二者结合,使 EDB 既能捕获详细的边缘信息,又能获得更抽象的全局特征,从而帮助模型更好地理解图像内容.最后,使用逐点卷积将通道数调整到所需维度.

## 1.2 DRRC

### 1.2.1 DRRC 整体框架

交通标志具有多样的形状、大小和颜色特征,不同类型的标志往往存在显著的尺度差异.为更好地捕捉不同尺度和不同复杂度的交通标志特征,设计 DRRC. DRRC 通过在 C2f 中集成 DRR 完成更丰富的卷积操作捕捉更多的特征信息,能够在多个尺度上提

取并融合特征.这使得模型不仅能够捕捉到小尺度的精细特征,还能保持对大尺度全局信息的敏感性. DRRC 数学上可表达为

$$F_{CO} = C_1 \left( \left( \frac{1}{2} C_1(F_{CI}) \right) \oplus \left( f_{DRR_N} \left( \frac{1}{2} C_1(F_{CI}) \right) \right) \right), \quad (3)$$

式中, $F_{CI}$  为 DRRC 的输入特征图, $F_{CO}$  为 DRRC 输出特征图, $C_1$  为  $1 \times 1$  大小的卷积, $f_{DRR_N}$  为经过  $N$  个 DRR 模块输出的  $N$  次特征图. DRRC 结构如图 3 所示.

### 1.2.2 基于 DRB 改进 DWR 的 DRR

DRR 通过将 DWR<sup>[9]</sup> 中的单一扩张卷积用 DRB<sup>[10]</sup>

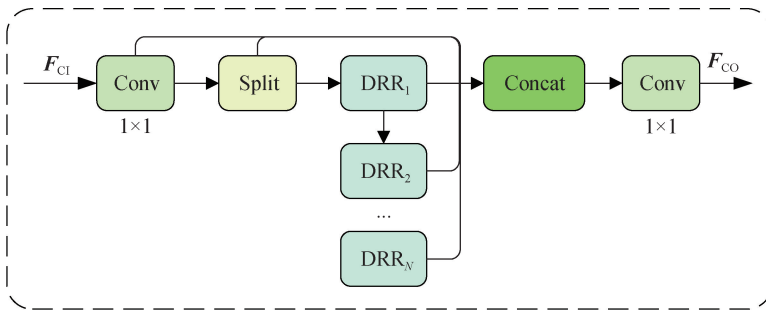


图 3 DRRC 结构

Fig. 3 DRRC structure

取代得到,DRR 与 DWR 结构如图 4 所示. DWR 是一种高效的两步多尺度上下文信息获取方法的模块,通常将区域特征图分成若干组,然后在不同组上进行不同扩张率的深度扩张卷积. DRR 将 DWR 中的两个深度扩张卷积替换为不同核尺寸的 DRB,将原本费力获取尽可能多的复杂语义信息的过程转变为简单地对每个简洁表达的特征图进行形态学滤波的过程. DRR 数学上可表达为:

$$F_2 = \frac{1}{3}C_3(F_{RI}), \tag{4}$$

$$F_{RO} = F_{RI} + C_1(f_{DC_3}(F_2) + f_{DRB_5}(F_2) + f_{DRB_7}(F_2)), \tag{5}$$

式中,  $F_{RI}$  为 DRR 输入特征图,  $C_3$  为  $3 \times 3$  大小的卷积,  $F_{RO}$  为 DRR 输出特征图,  $F_2$  为经过卷积处理的初级特征图,  $f_{DC_3}$  为  $3 \times 3$  大小的深度卷积操作,  $f_{DRB_k}$  为  $k \times k$  大小的 DRB 操作.

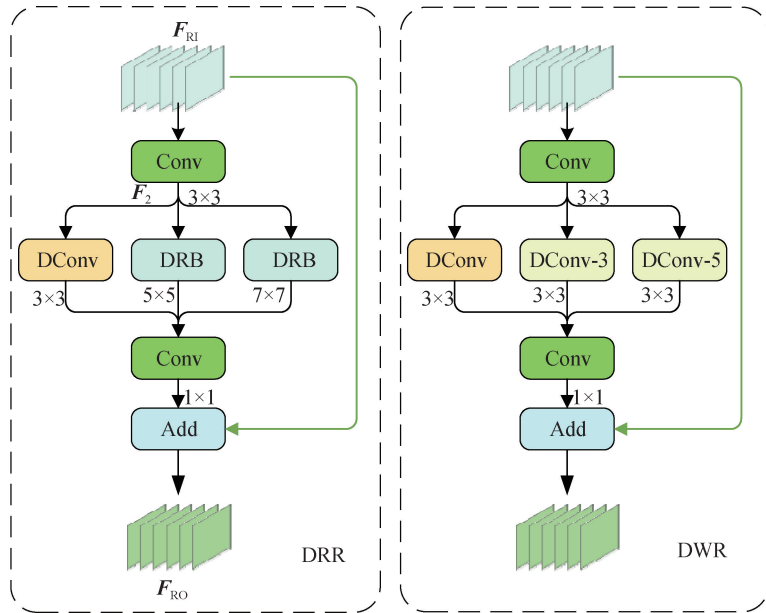


图 4 DRR 和 DWR 的结构

Fig. 4 Structures of DRR and DWR

### 1.2.3 DRB

DRB 使用一个非扩展的小核和多个扩展的小核层来增强一个非扩展的大核转换层. 通过增强大核捕捉稀疏模式的能力从而产生更高质量的特征. 为了消除额外扩张卷积层的推理成本,将整个块等效地转换为一个单一的非扩张卷积层进行推理. DRB 结构如图 5 所示. DRB 数学上可表达为

$$F_{BK} = C_K(F_{BI}) + \sum_{i=1}^K C_{ki,di}^{DI}(F_{BI}), \tag{6}$$

式中,  $F_{BI}$  为 DRB 输入特征图,  $F_{BK}$  为 DRB 输出特征图,  $C_K$  为  $K \times K$  大小的大核卷积,  $C_{ki,di}^{DI}$  为第  $i$  个大小为  $k \times k$ 、膨胀率为  $d$  的扩张卷积.

DRR 引入多尺度扩张卷积,旨在捕捉不同尺度特征、扩展感受野和捕捉更多的上下文信息. 与单一扩

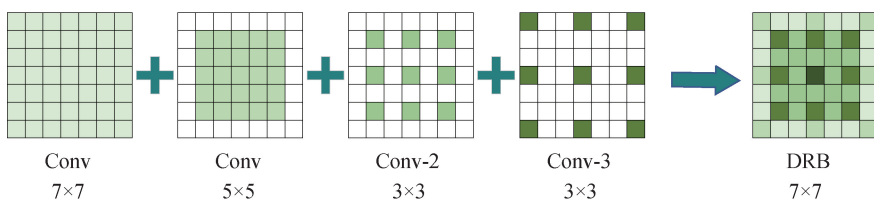


图 5 DRB 结构

Fig. 5 Structure of DRB

张率的深度扩张卷积相比,通过多种卷积核尺寸和扩张率的组合,能够实现更高效的多尺度特征融合,有助于模型同时捕捉到细节信息和全局上下文。

相较于 C2f, DRRC 可以在更多尺度上进行特征提取和融合,使得模型能够更加准确地识别和分类各种尺寸的交通标志,可以增强模型应对复杂背景的能力。使用不同扩张率的卷积核能够同时提取到细节和全局信息,增强了特征图的表达能力,对于处理不同距离、大小和形状的交通标志较为重要。

### 1.3 RCG-FPN

为深度融合浅层与深层特征信息,引导模型对每个通道重要信息的关注,设计 RCG-FPN,结构如图 1 虚线框所示。

在卷积神经网络中,靠近输入的浅层特征图由于经过的卷积操作较少,从而具有较高的分辨率和较小的感受野,有效保留了局部细节信息。深层特征图则具有丰富的语义信息和较大的感受野,有利于大目标的识别。但因经过多次卷积操作和采样处理,深层特征图的分辨率降低的同时,通道数也会显著增加,而较高的通道数则意味着卷积核在操作时需要与更多通道进行计算,导致参数量增大。由于交通标志在图

像中通常占据较小的像素区域,较高分辨率的特征图具有较为丰富的边界、位置和纹理等细节特征,因此,RCG-FPN 选择增加微小目标检测层,引入浅层语义信息来增强对复杂背景中小目标交通标志的检测,同时通过移除大目标检测层来减少较多的参数量,提升模型的轻量化程度。

#### 1.3.1 SBA

YOLOv8 在使用简单的拼接操作处理尺寸差异较大的目标时,可能会导致某些重要的特征被削弱或丢失,因此引入 SBA<sup>[11]</sup> 改进拼接操作同时完成采样功能。SBA 选择性地聚合边界信息和语义信息,以描绘更细粒度的目标轮廓和重新校准目标的位置。通过重新校准注意单元(recalibration attention unit, RAU) 块,在融合前自适应地从两个待融合的层级特征 ( $F^s$ ,  $F^b$ ) 中提取相互表示,并将浅层和深层信息以不同的方式馈送到两个 RAU 块中,以弥补高层语义特征的空间边界信息缺失和低层语义特征的语义信息缺失。最后,使用  $3 \times 3$  卷积将两个 RAU 块的输出连接起来。这种聚合策略实现了不同特征的鲁棒组合,并细化了粗糙特征。SBA 与 RAU 如图 6 所示。

RAU 块的核心计算过程  $f_{RAU}$  可表示如下:

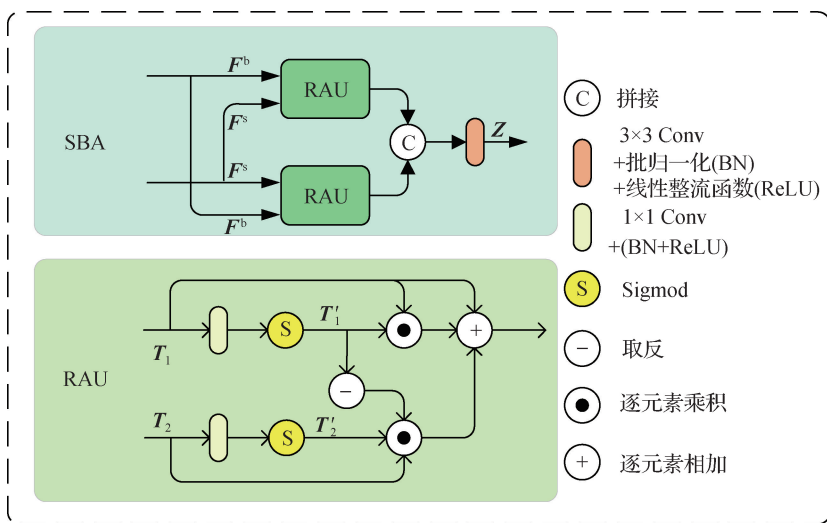


图 6 SBA 和 RAU 结构

Fig. 6 Structure of SBA and RAU

$$\mathbf{T}'_1 = W_\theta(\mathbf{T}_1), \mathbf{T}'_2 = W_\phi(\mathbf{T}_2), \quad (7)$$

$$f_{\text{RAU}}(\mathbf{T}_1, \mathbf{T}_2) = \mathbf{T}'_1 \odot \mathbf{T}_1 + \mathbf{T}'_2 \odot \mathbf{T}_2 \odot R(\mathbf{T}'_1) + \mathbf{T}_1, \quad (8)$$

式中:  $\mathbf{T}_1$  和  $\mathbf{T}_2$  为输入特征;  $W$  为如图 6 所示的两个线性映射和 Sigmoid 函数操作, 将通道维度减少至 32, 并获得特征图  $\mathbf{T}'_1$  和  $\mathbf{T}'_2$ ; 符号“ $\odot$ ”表示逐元素乘法;  $R$  为反向操作,  $R(\mathbf{T}'_1) = 1 - \mathbf{T}'_1 / \|\mathbf{T}'_1\|$ , 将不精确和粗略的估计优化为更加准确和完整的预测图。

因此, SBA 模块的计算过程可表示为

$$\mathbf{Z} = f_{c_3}(f_{\text{RAU}}(\mathbf{F}^s, \mathbf{F}^b) \oplus f_{\text{RAU}}(\mathbf{F}^b, \mathbf{F}^s)), \quad (9)$$

式中:  $f_{c_3}$  表示带有批归一化和 ReLU 激活层的  $3 \times 3$  卷积,  $\mathbf{F}^s \in \mathbb{R}^{\frac{H}{8} \times \frac{W}{8} \times 32}$  包含深层语义信息,  $\mathbf{F}^b \in \mathbb{R}^{\frac{H}{4} \times \frac{W}{4} \times 32}$  包含丰富的边界细节信息,  $\mathbf{Z} \in \mathbb{R}^{\frac{H}{4} \times \frac{W}{4} \times 32}$  是 SBA 模块的输出特征图。

SBA 通过引入自适应权重来控制不同尺度特征的融合程度, 能够更有效地结合来自不同尺度的信息, 通过注意力机制优化特征图的表达, 确保在每个

层级上提取到的特征都具有更高的区分能力, 使得较小或较大的目标在融合过程中不被忽略。

### 1.3.2 CGAFusion

在交通标志检测任务中, 树木、建筑物、车辆等复杂背景可能干扰模型对交通标志的识别. 传统的 FPN 结构缺乏对这些复杂背景的处理, 可能导致模型对背景噪声的过度敏感; 又因浅层特征信息与深层特征信息具有完全不同的感受野, 简单的加法或拼接操作无法解决融合前的感受野失配问题. 故本文引入 CGAFusion<sup>[12]</sup> 使模型专注于图像中最重要的部分, 减少对不相关区域的注意力. CGAFusion 将低级和相应的高级特征输入到内容引导注意力 (content-guided attention, CGA) 机制中计算不同信息的权重以调制特征, 采用加权求和的方法进行组合以自适应地融合特征. 同时, 通过跳跃连接添加输入特征, 以缓解梯度消失问题和简化学习过程. CGAFusion 和 CGA 详细过程如图 7 所示。

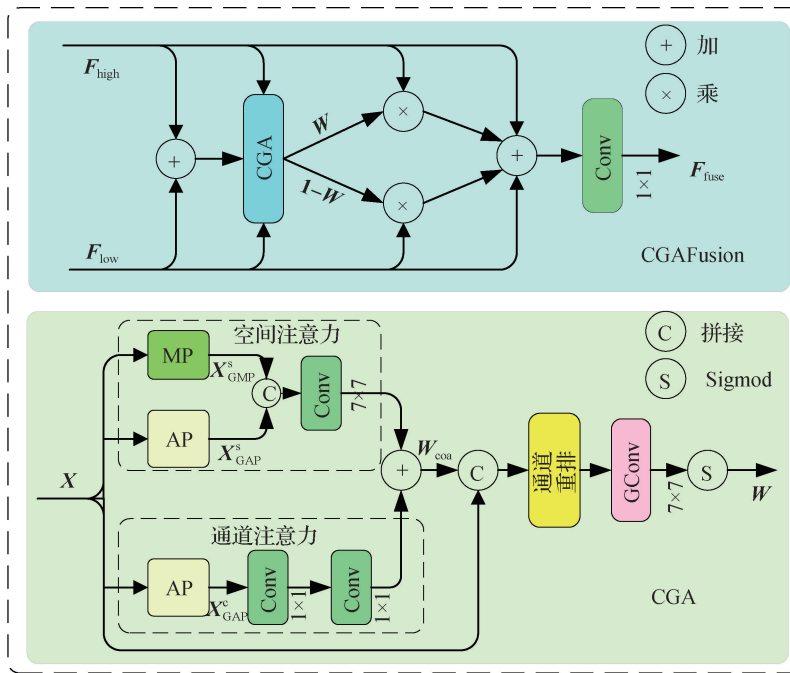


图 7 CGAFusion 与 CGA 结构

Fig. 7 Structures of CGAFusion and CGA

CGA 的运算过程如下:

$$W_{\text{coa}} = C_1(\max(0, C_1(X_{\text{GAP}}^c))) + C_7(X_{\text{GAP}}^s \oplus X_{\text{GMP}}^s), \quad (10)$$

式中,  $\max(0, x)$  表示 ReLU 激活函数,  $X_{\text{GAP}}^c$ ,  $X_{\text{GAP}}^s$  和  $X_{\text{GMP}}^s$  分别表示经过全局平均池化操作在通道维度、全局平均池化操作在空间维度、以及全局最大池化操作在空间维度处理的特征。

为了获得最终精细化的空间重要性映射  $W$ , 根据输入特征对  $W_{\text{coa}}$  的每个通道进行调整。

$$W = \sigma(C_7^c(f_{\text{CS}}(X \oplus W_{\text{coa}}))), \quad (11)$$

式中,  $\sigma$  为 Sigmoid 激活函数,  $f_{\text{CS}}$  为通道洗牌操作,  $C_k^c$  为核大小  $k \times k$  的分组卷积层 (GConv)。

如图 7 所示, CGAFusion 可以表达为

$$F_{\text{fuse}} = C_1(F_{\text{low}} \cdot W + F_{\text{high}} \cdot (1 - W) +$$

$$\mathbf{F}_{\text{low}} + \mathbf{F}_{\text{high}}), \quad (12)$$

式中:  $\mathbf{F}_{\text{low}}$  为输入的低级特征,  $\mathbf{F}_{\text{high}}$  为输入的高级特征,  $\mathbf{F}_{\text{fuse}}$  为最终的融合特征.

综上所述,CGA 以粗到精的方式获得每一条输入特征通道的专属空间重要性图(spatial importance map, SIM). CGA 首先通过融合空间注意和通道注意,得到粗 SIM. 粗 SIM 和输入特征的每个通道都通过通道洗牌操作以交替的方式重新排列,结合后续的分组卷积层,可以减少参数的数量. CGA 为每个通道分配一个唯一的 SIM,以引导模型关注每个通道的重要区域. CGAFusion 利用 CGA 计算的空间和通道权重来调制特征,并通过加权求和的方法自适应地融合编码器部分的低级特征与相应的高级特征.

### 1.4 Wise-SIoU 损失函数

考虑到锚框与真实框存在不匹配的问题,以及标注数据的锚框质量有好有坏,为了提高锚框与真实框之间的匹配程度,引导模型更关注普通质量的锚框,融合 Wise-IoU<sup>[13]</sup> 与 SIoU<sup>[14]</sup> 的思想,设计损失函数 Wise-SIoU.

#### 1.4.1 SIoU

为提高锚框与真实框之间的匹配程度,以及加快锚框回归速度,SIoU 通过对角度成本、距离成本和形状成本三个方面进行不匹配惩罚.

若  $\Delta$ 、 $\Delta$  和  $\Omega$  分别为角度成本、距离成本和形状成本,则 SIoU 损失函数可表示为:

$$\mathcal{L}_{\text{SIoU}} = \mathcal{L}_{\text{IoU}} + \frac{\Delta + \Omega}{2}, \quad (13)$$

$$\Delta = \sum_{t=x,y} (1 - e^{-(2-\Delta)t_t}), \rho_x = \left(\frac{c_w}{W_g}\right)^2, \rho_y = \left(\frac{c_g}{H_g}\right)^2, \quad (14)$$

$$\Delta = \sin(2\arcsin(x)), x = \frac{c_g}{\sigma} = \sin(\alpha), \quad (15)$$

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega_t})^4, \omega_w = \frac{|\tau - \tau^{\text{gt}}|}{\max(\tau, \tau^{\text{gt}})}, \omega_h = \frac{|h - h^{\text{gt}}|}{\max(h, h^{\text{gt}})}, \quad (16)$$

式中:  $c_w$ 、 $c_g$ 、 $\sigma$  以及  $\alpha$  表示锚框与真实框的中心点在  $x$  方向上的距离、 $y$  方向上的距离、欧式距离以及两者连线与  $x$  方向上的夹角.  $W_g$ 、 $H_g$  表示锚框与真实框最小包围框的宽和高,  $\tau$ 、 $h$  表示锚框的宽和高,  $\tau^{\text{gt}}$ 、 $h^{\text{gt}}$  表示真实框的宽和高. 式中各参数如图 8 所示.

#### 1.4.2 Wise-IoU

Wise-IoU 是一种基于交并比(IoU)的动态非单调聚焦损失,其采用离群度评价锚框的质量,提供一

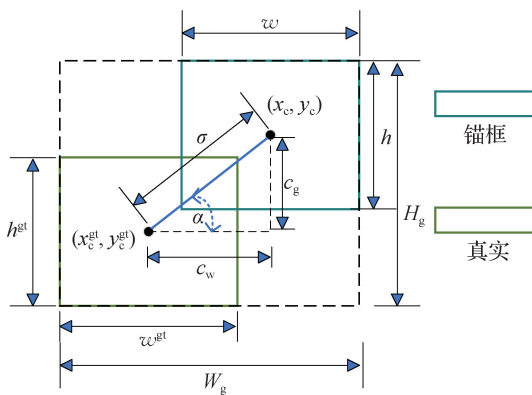


图 8 SIoU 各参数示意

Fig. 8 Parameter illustration of Wise-SIoU

种明智的梯度增益分配策略. 这种策略降低了高质量锚框的竞争力,同时也减少了低质量样本产生的有害梯度. 数学上表达如下:

$$\mathcal{L}_{\text{WIoU}} = r \mathcal{R}_{\text{WIoU}} \mathcal{L}_{\text{IoU}}, \quad (17)$$

式中:  $\mathcal{R}_{\text{WIoU}} \in [1, e)$  将显著放大普通质量锚框的  $\mathcal{L}_{\text{IoU}}$ .  $r$  为由离群度  $\beta$  构造的非单调聚焦系数,离群度  $\beta$  越小意味着锚框质量越高. 数学上表达如下:

$$r = \frac{\beta}{\delta \tau^{\beta-\delta}}, \beta = \frac{\mathcal{L}_{\text{IoU}}^*}{\mathcal{L}_{\text{IoU}}} \in [0, +\infty), \quad (18)$$

$$\mathcal{R}_{\text{WIoU}} = \exp\left(\frac{\sigma^2}{(W_g^2 + H_g^2)^*}\right), \quad (19)$$

式中:  $\delta$  和  $\tau$  为超参数; \* 表示分离操作,在反向传播时,不对被 \* 标记的变量求导,仅作为离群度的参考基准,以此有效地排除它们参与反向传播. 动态的锚框质量划分标准,使得在每一个时刻都能做出最符合当前情况的梯度增益分配策略.

#### 1.4.3 Wise-SIoU

综合 SIoU 与 Wise-IoU 的思想,Wise-SIoU 可以表示为

$$\mathcal{L}_{\text{Wise-SIoU}} = r \mathcal{R}_{\text{WIoU}} \mathcal{L}_{\text{SIoU}}. \quad (20)$$

## 2 实验结果

### 2.1 实验数据与环境

为验证 EDR-YOLOv8 的先进性与泛化性,在 TT100K 与 CCTSDB 2021 数据集上进行实验. TT100K 数据集由清华与腾讯联合实验室整理并公布,共计 10 万张图像,包含 30 000 个交通标志实例,221 种不同类别,其中 128 类有标注,类别标注如: i10-右转、p5-禁止掉头和 w9-注意落石等. TT100K 原始数据集中各个类别的数量存在较大差异,因此对数据

集进行预处理,筛选出具有 100 个以上实例的 42 个类别,共计 6 866 张图像作为本次实验数据集,按 7 : 3 随机划分训练集与测试集. CCTSDB 2021 数据集由长沙理工大学的相关学者及团队制作而成,其有交通标志样本图片近 20 000 张,共含交通标志近 40 000 个,同时包含按尺寸、场景分类区分的图像数据集,含有

丰富的道路背景信息. 将交通标志按含义分为指示、禁令和警告标志三类. 实验按照 7 : 3 划分训练集与测试集,从其训练集中随机抽取 3 500 张用作本次实验训练集,测试集选用原数据集中已经划分的 1 500 张测试集. TT100K 数据集与 CCTSDB 2021 数据集对交通标志的标注方式如图 9 所示.

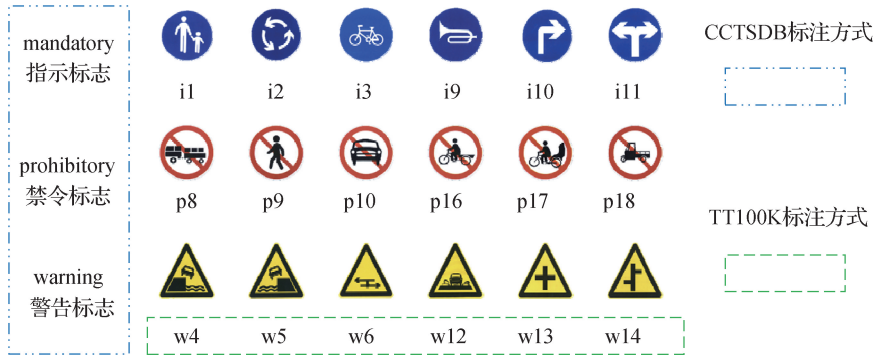


图 9 交通标志类别划分  
Fig. 9 Classification of traffic signs

实验在操作系统 Windows11 下运行, GPU 为 NVIDIA RTX 3060, 内存为 16 GB, 使用编程语言 Python 3. 8, 使用 cuda 11. 7 对 GPU 进行加速, 深度学习框架为 pytorch 1. 17. 1. 实验参数设置为: 优化器选用 SGD, 批大小为 16, 图像大小为 640 × 640 像素, 进程数为 8, 在最后 10 轮关闭数据增强, 学习率为 0. 01, 动量因子为 0. 937, 训练轮次为 300.

实验使用精确率  $P$ 、平均精度均值  $P_{mA}$ 、参数量  $N_{params}$ 、浮点运算速度  $V$ 、 $F_1$  分数和模型尺寸  $S$  对模型进行评价.  $P$  是模型正确预测为正例的样本数量占预测为正例样本数量的比重. 召回率  $R$  指模型正确预测为正例的样本数量占所有真正为正例的样本数量的比例.  $P$  和  $R$  计算公式如下:

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}, R = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (21)$$

式中:  $N_{TP}$  为真正例, 表示预测为正例实际也为正例的样本数量;  $N_{FP}$  为假正例, 表示预测为正例实际为反例的样本数量,  $N_{FN}$  为假反例, 指预测为反例实际为正例的样本数量.

平均精确度均值  $P_A$  指预测图片内某一类别  $P(R)$  曲线下的面积, 表示在一组不同 IoU 阈值下某一类别的平均准确度.  $P_{mA}$  则指  $N$  类样本在某一 IoU 阈值时所有种类样本平均准确率  $P_A$  的平均值, 计算式如下:

$$P_{mA@0.5} = \frac{1}{N} \sum_{i=1}^N P_{A@0.5}^i, P_A = \int_0^1 P(R) dR, \quad (22)$$

$$P_{mA@[0.50,0.95]} = \frac{1}{10} (P_{mA@0.50} + P_{mA@0.55} + \dots + P_{mA@0.95}), \quad (23)$$

其中, 下标 @ 后的取值表示 IoU 阈值的取值范围.  $F_1$  分数是综合考虑精确度和召回率的指标,  $F_1$  值越大代表模型的性能越好, 计算公式如下:

$$F_1 = \frac{2PR}{P+R}. \quad (24)$$

## 2.2 消融与对比实验

### 2.2.1 消融实验

为了验证本文改进模型的有效性, 以原始模型 YOLOv8n 为基线模型, 在 TT100K 数据集上对各个改进部分进行消融实验, 结果如表 1 所示.

特征提取模块 EDB 相对于 CBS,  $P$ 、 $P_{mA@0.50}$ 、 $P_{mA@[0.50,0.95]}$  和  $F_1$  分别提升 1.3、1.7、0.8 和 7.7 个百分点, 说明重点关注边缘信息的 EDB 模块更适合交通标志的识别, 这主要与交通标志由对比度高的线条构成有关. 基于 C2f 改进的多尺度感知模块 DRRC 相对于 C2f,  $P$ 、 $P_{mA@0.50}$  和  $F_1$  分别提升 3.4、0.5 和 1.9 个百分点,  $N_{params}$  减少  $10^5$ ,  $v$  提升  $0.2 \times 10^9/s$ , 说明 DRRC 在提升信息感知能力的同时, 在降低参数量和计算复杂度方面均有改善; RCG-FPN 深度融合交通标志的低级与高级特征, 虽然计算复杂度明显增加, 但模型的参数量降低了  $0.6 \times 10^6$ ,  $P_{mA@0.50}$ 、 $P_{mA@[0.50,0.95]}$  和  $F_1$  也分别提升 1.9、2.6 和 1.9 个百分点; Wise-SIoU 具有明智的梯度增益分配策略, 在不改变参数量和计

表 1 EDR-YOLOv8 在 TT100K 数据集上的消融实验结果  
Tab. 1 The ablation experiment results of EDR-YOLOv8 on the TT100K dataset

YOLOv8	EDB	DRRC	RCG-FPN	Wise-SIoU	$P/\%$	$P_{m\text{A}@0.50}/\%$	$P_{m\text{A}@[0.50,0.95]}/\%$	$F_1/\%$	$N_{\text{params}}/10^6$	$v/(10^9 \cdot \text{s}^{-1})$
✓	—	—	—	—	73.6	73.1	55.3	63.2	3.0	<b>8.1</b>
✓	✓	—	—	—	74.9	74.8	56.1	70.9	3.0	8.3
✓	✓	✓	—	—	78.3	75.3	55.9	72.8	2.9	<b>8.1</b>
✓	✓	✓	✓	—	78.5	77.2	58.5	74.7	2.3	14.9
✓	✓	✓	✓	✓	<b>82.8</b>	<b>79.6</b>	<b>61.0</b>	<b>75.3</b>	<b>2.3</b>	14.9

算复杂度的前提下,  $P$ 、 $P_{m\text{A}@0.50}$ 、 $P_{m\text{A}@[0.50,0.95]}$  和  $F_1$  分别提升 4.3、2.4、2.5 和 0.6 个百分点。

由表 1 可以看出, 本文提出的 EDR-YOLOv8 相比原模型, 在  $P$ 、 $P_{m\text{A}@0.50}$ 、 $P_{m\text{A}@[0.50,0.95]}$  和  $F_1$  分别提升 9.2、6.5、5.7 和 12.1 个百分点, 提升率分别为 12.5%、8.9%、10.3% 和 19.1%, 模型参数量也下降了 23%, 证明了本文改进模型在硬件要求上更高, 但在轻量化和识别性能上更有优势。

### 2.2.2 对比实验

为验证模型的优越性, 在 TT100K 数据集上对本文改进模型与当前主流模型进行对比实验, 结果如表

2 所示。在轻量化方面, EDR-YOLOv8 与文献[8]具有最小规模的参数  $2.3 \times 10^6$ , 仅为 RT-DETR 的 1/12 不到。模型大小仅次于文献[8], 但模型性能优于文献[8],  $P$ 、 $P_{m\text{A}@0.50}$ 、 $P_{m\text{A}@[0.50,0.95]}$ 、 $F_1$  相对于文献[8]分别高出 6.8、3.6、4.3、2.6 个百分点。在模型性能方面, EDR-YOLOv8 的  $P$ 、 $P_{m\text{A}@0.50}$ 、 $F_1$  均取得最优值, 相对于文献[20]的次优结果分别高出 1.9、0.2、0.7 个百分点。 $P_{m\text{A}@[0.50,0.95]}$  虽略低于文献[20]0.7 个百分点, 但算法复杂度仅为文献[20]的 61%。在算法复杂度方面, EDR-YOLOv8 虽不具有突出表现, 但也具有中等表现, 约为复杂度最低算法 YOLOv12n 的 2.5 倍, 复杂度最高算法 RT-DETR 的 1/7。

表 2 不同模型在 TT100K 数据集上的对比实验结果  
Tab. 2 The comparative experimental results of different models on the TT100K dataset

模型	$P/\%$	$P_{m\text{A}@0.50}/\%$	$P_{m\text{A}@[0.50,0.95]}/\%$	$F_1/\%$	$N_{\text{params}}/10^6$	$v/(10^9 \cdot \text{s}^{-1})$	S/MB
YOLOv5s	69.4	65.9	48.9	65.5	7.1	16.1	14.0
YOLOv6n	63.6	64.0	47.8	61.7	4.6	11.4	10.0
YOLOv7-tiny	54.4	50.7	35.9	54.9	6.1	13.5	11.9
YOLOv8n	73.6	73.1	55.3	63.2	3.0	<b>8.1</b>	6.0
YOLOv10n <sup>[15]</sup>	75.2	69.8	52.9	67.5	2.7	8.3	5.6
YOLOv11n <sup>[16]</sup>	69.8	68.7	52.4	66.6	2.6	6.4	5.3
YOLOv12n <sup>[17]</sup>	56.6	59.1	44.3	56.9	2.5	5.9	5.3
RT-DETR <sup>[18]</sup>	67.2	62.0	48.7	62.3	28.53	100.8	56.4
文献[8]	76.0	76.0	56.7	72.7	<b>2.3</b>	11.0	<b>4.8</b>
文献[19]	75.7	76.1	58.2	72.7	2.5	12.1	5.2
文献[20]	80.9	79.4	<b>61.7</b>	74.6	2.4	24.5	5.0
文献[21]	76.9	77.3	58.0	73.0	2.4	23.6	5.3
EDR-YOLOv8	<b>82.8</b>	<b>79.6</b>	61.0	<b>75.3</b>	<b>2.3</b>	14.9	5.0

为进一步验证改进模型的泛化性能, 在 CCTSDB 数据集上对本文改进模型与当前主流模型进行对比实验, 结果如表 3 所示。

同样地, 在 CCTSDB 上, EDR-YOLOv8 除在计算

速度上有所欠缺外, 其他各项指标均表现最佳或接近最佳, 在轻量化和识别性能方面表现优异。结果表明 EDR-YOLOv8 在其他数据集上也有较好的表现, 具有较好的泛化性。

表 3 不同模型在 CCTSDB 数据集上的对比实验结果

Tab. 3 The comparative experimental results of different models on the CCTSDB dataset

模型	$P/\%$	$P_{m\text{A}@0.50}/\%$	$P_{m\text{A}@[0.50,0.95]}/\%$	$F_1/\%$	$N_{\text{params}}/\text{M}$	$v/(10^9 \cdot \text{s}^{-1})$	$S/\text{MB}$
YOLOv5s	82.4	67.7	39.9	69.4	7.0	15.8	13.7
YOLOv6n	82.3	69.6	40.9	72.7	4.6	11.3	10.0
YOLOv7-tiny	69.0	53.6	27.7	56.2	6.0	13.2	11.7
YOLOv8n	83.0	69.6	42.9	70.9	3.0	8.1	6.0
YOLOv10n	<b>84.0</b>	69.9	42.2	71.1	2.7	8.2	5.5
YOLOv11n	81.5	71.0	43.7	72.3	2.6	6.3	5.2
YOLOv12n	82.8	69.1	41.1	70.7	2.5	<b>5.8</b>	5.2
RT-DETR	78.3	72.0	42.1	71.6	28.5	100.6	56.3
文献[8]	82.7	72.9	44.6	72.4	2.2	10.2	<b>4.6</b>
文献[19]	83.2	74.5	45.1	74.2	2.4	11.3	4.9
文献[20]	83.1	73.0	43.9	72.3	2.4	24.3	4.9
文献[21]	82.9	74.2	<b>46.0</b>	74.1	2.4	23.4	4.9
EDR-YOLOv8	83.8	<b>74.6</b>	45.6	<b>74.5</b>	2.3	14.1	4.7

### 2.3 颈部网络和损失函数相关的对比实验

为验证 RCG-FPN 结构的有效性,在 TT100K 数据集上进行改进实验,其中“+SD”、“-BD”、“+CGAFusion”和“+SBA”分别指在 YOLOv8 中依次进行加入微小目标检测层、移除大目标检测层、加入 CGAFusion 模块和引入 SBA 模块的操作,结果如表 4 所示。

由表 4 可以看出,增加微小目标检测的 YOLOv8,虽然计算量增长,但各个指标相较原 YOLOv8 算法均有较大提升;通过移除大目标检测层,模型参数量得以显著减少,但伴随少量检测性能的下降;在此基础上进一步构建 RCG-FPN 结构,模型在保持较低参数量的前提下实现了检测效果的进一步提升。

表 4 RCG-FPN 在 TT100K 数据集上的改进实验

Tab. 4 Improved experiment of RCG-FPN on TT100K dataset

模型	$P/\%$	$P_{m\text{A}@0.50}/\%$	$P_{m\text{A}@[0.50,0.95]}/\%$	$F_1/\%$	$N_{\text{params}}/10^6$	$v/(10^9 \cdot \text{s}^{-1})$
YOLOv8	73.6	73.1	55.3	63.2	3.0	<b>8.1</b>
+SD	78.5	<b>78.1</b>	59.2	73.3	3.0	13.1
-BD	76.0	77.4	58.6	70.4	<b>2.1</b>	12.4
+CGAFusion	<b>78.8</b>	77	59.0	72.7	2.1	12.7
+SBA	78.7	77.8	<b>59.3</b>	<b>73.6</b>	2.2	14.7

为继续验证 RCG-FPN 的优越性,在 YOLOv8 上融入其他主流 Neck 结构:GFPN<sup>[22]</sup>、BiFPN<sup>[23]</sup>、SlimNeck<sup>[24]</sup>、AFPN<sup>[25]</sup> 在 TT100K 数据集上进行对比试验,结果如表 5 所示。

从表 5 可以看出,本文所提出 RCG-FPN 由于引入了小目标检测层导致计算量较高,但在  $P$ 、 $P_{m\text{A}@0.50}$ 、 $P_{m\text{A}@[0.50,0.95]}$  及  $F_1$  上均为最优结果,且参数量处于最低水平,具有相当不俗的表现。RCG-FPN 相较其他

Neck 结构,在提升 YOLOv8 算法识别性能和实现轻量化方面具有一定的优越性。

为验证 Wise-SIoU 损失函数的优越性,在 TT100K 数据集上,对 EDR-YOLOv8 模型分别采用 Ciou、Shape-IoU<sup>[26]</sup>、Wise-IoU、Siou、Inner-IoU<sup>[27]</sup>、Wise-SIoU 损失函数进行训练,训练结果如图 10 所示,其中 EDR 表示 EDR-YOLOv8。

如图 10 所示,Wise-IoU 的梯度增益分配策略有效

表 5 不同 Neck 结构的模型在 TT100K 数据集上的对比实验结果

Tab. 5 The comparative experimental results of models with different Neck structures on the TT100K dataset

模型	$P/\%$	$P_{m\Lambda@0.50}/\%$	$P_{m\Lambda@[0.50,0.95]}/\%$	$F_1/\%$	$N_{Params}/10^6$	$v/(10^9 \cdot s^{-1})$
YOLOv8+GFPN	75.5	72.1	54.3	69.6	3.3	8.3
YOLOv8+BiFPN	72.2	73.5	55.9	69.9	<b>2.0</b>	<b>7.2</b>
YOLOv8+SlimNeck	72.9	72.7	55.4	70.5	2.8	7.3
YOLOv8+AFPN	64.9	64.3	48.3	62.5	2.6	8.6
YOLOv8+RCG-FPN	<b>78.7</b>	<b>77.8</b>	<b>59.3</b>	<b>74.1</b>	2.2	14.7

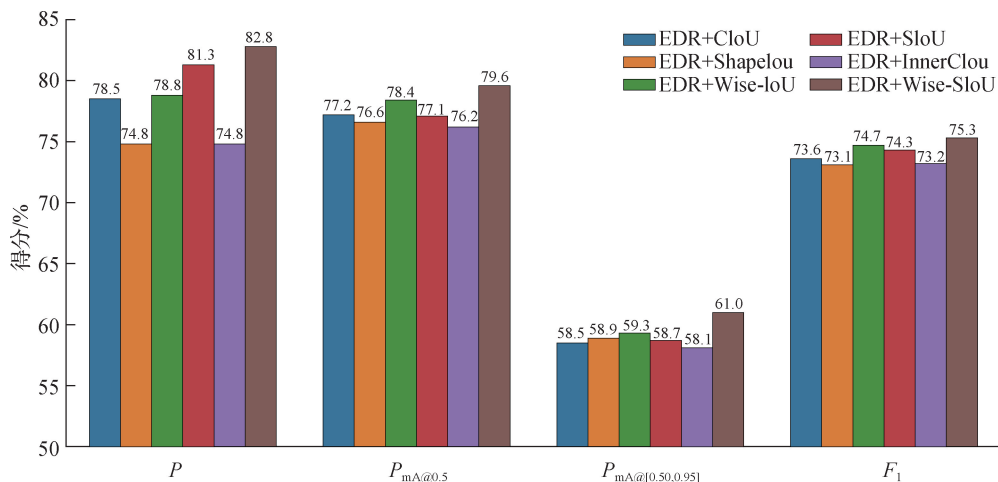


图 10 改进模型在不同损失函数下在 TT100K 数据集上的对比实验结果

Fig. 10 The comparative experimental results of the improved model under different loss functions on the TT100K dataset

地改善了检测效果,在  $P$  和  $P_{m\Lambda}$  上均带来了不同程度的提升;SIoU 对于不匹配方向的惩罚较大提升了模型的准确率;Wise-SIoU 则综合了 Wise-IoU 和 SIoU 的优势,取得最优的检测效果.

### 2.4 实验效果

为验证 EDR-YOLOv8 的改进效果,在同一环境下,对同一组含交通标志的图像进行 EDR-YOLOv8 与其他交通标志算法的对比实验,结果如图 11 所示,图中不同的标签用不同颜色表示.

图 11 中的图像已对交通标志部分进行截取放大展示.对于图像(a)的交通标志,YOLOv8 与文献[21]算法均未能正确识别,文献[20]算法虽有正确识别,但置信度低于本文算法;对于图像(b)中的交通标志,YOLOv8 算法错检标签为“w57”的交通标志,文献[20]与文献[21]的算法漏检标签为“w13”的交通标志,本文算法全部正确检测;对于图像(c)的交通标志,文献[20]算法与文献[21]算法均错检两个交通标志,YOLOv8 与本文算法虽漏检一个交通标志,但另外两个交通标志均识别正确.综合来看,本文算法相较其

他算法对交通标志的检测效果更好.

为验证本文算法在真实环境中的检测效果,用基于 TT100K 数据集训练的本文算法对真实拍摄的含交通标志的图像进行检测,结果如图 12 所示.图中检测到的交通标志的标注类别均符合 TT100K 数据集中划分的对应类别,且均正确识别.由此可见,对于现实场景中常见的指示、禁令和警告类交通标志,本文算法仍有较好的识别效果.

### 3 结束语

针对交通标志检测过程中易出现漏检、误检的问题,本文提出了一种空间信息聚合网络模型 EDR-YOLOv8.EDB 侧重边缘信息检测,更适合由对比度高的线条构成的交通标志的识别;设计具有不同感受野的多尺度模块 DRRC 高效地获取上下文语义信息,能更好地区分背景信息与交通标志;RCG-FPN 聚合局部与全局信息,强化低级与高级特征的有效融合,虽增加了计算复杂度,但降低了算法参数量,并提升了识别准确率;Wise-SIoU 能很好地减少低质量与高

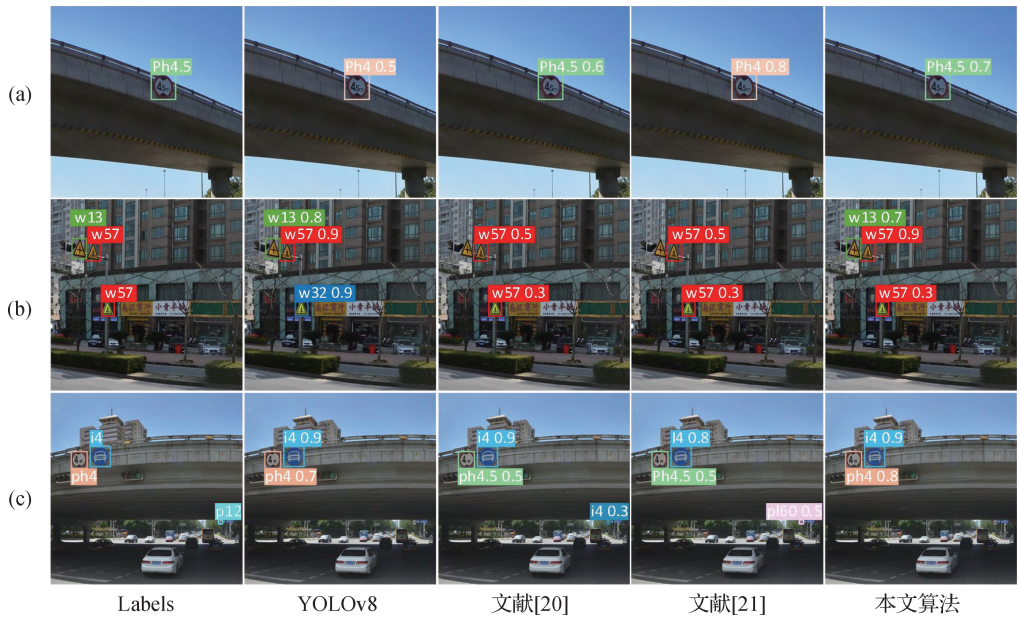


图 11 EDR-YOLOv8 与其他交通标志算法在 TT100K 数据集上的检测效果对比

Fig. 11 Comparison of the detection effect of EDR-YOLOv8 with other traffic sign algorithms on the TT100K dataset



图 12 EDR-YOLOv8 真实场景检测效果

Fig. 12 EDR-YOLOv8 detection effect on real-world scene

质量锚框对模型的影响,并加快锚框的回归.改进后的 YOLOv8 网络模型在 TT100K 数据集和 CCTSDB 2021 数据集上对交通标志的检测效果在各项指标上均优于原模型,具有精度高、参数量少的优点,相比其他模型也具有不同程度的优越性.

参考文献:

[1] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and

Machine Intelligence, 2017, 39(6): 1137-1149.  
 [2] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 386-397.  
 [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 779-788.  
 [4] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C] // Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 7263-7271.

- [5] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]// Computer Vision-ECCV. Cham: Springer, 2016: 21-37.
- [6] 井方科, 任红格, 李松. 基于多尺度特征融合的小目标交通标志检测[J]. 激光与光电子学进展, 2024, 61(12): 372-380.
- [7] 朱宁可, 葛青, 王翰文, 等. 基于 YOLOv5-MGC 的实时交通标志检测[J]. 激光与光电子学进展, 2024, 61(12): 348-357.
- [8] 高良鹏, 赵博文, 简文良. 基于 Faster-YOLOv8 网络模型的车载交通标志检测算法研究[J]. 重庆交通大学学报(自然科学版), 2024, 43(8): 114-123.
- [9] WEI H R, LIU X, XU S C, et al. DWRSeg: rethinking efficient acquisition of multi-scale contextual information for real-time semantic segmentation[EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2212.01173v3>.
- [10] DING X H, ZHANG Y Y, GE Y X, et al. UniRepLkNet: a universal perception large-kernel ConvNet for audio, video, point cloud, time-series and image recognition [EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2311.15599v2>.
- [11] TANG F L, XU Z X, HUANG Q M, et al. DuAT: dual-aggregation transformer network for medical image segmentation[C]// Pattern Recognition and Computer Vision. Cham: Springer, 2023: 343-356.
- [12] CHEN Z X, HE Z W, LU Z M. DEA-Net: single image dehazing based on detail-enhanced convolution and content-guided attention [J]. IEEE Transactions on Image Processing, 2024, 33: 1002-1015.
- [13] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2301.10051v3>.
- [14] GEVORGYAN Z. SiIoU loss: more powerful learning for bounding box regression[EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2205.12740v1>.
- [15] WANG A, CHEN H, LIU L H, et al. YOLOv10: real-time end-to-end object detection [EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2405.14458v2>.
- [16] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2410.17725v1>.
- [17] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: attention-centric real-time object detectors [EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2502.12524v1>.
- [18] ZHAO Y A, LV W Y, XU S L, et al. DETRs beat YOLOs on real-time object detection [C]// Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 16965-16974.
- [19] 熊恩杰, 张荣芬, 刘宇红, 等. 面向交通标志的 Ghost-YOLOv8 检测算法[J]. 计算机工程与应用, 2023, 59(20): 200-207.
- [20] 闫世洋, 罗素云. 基于 SC-YOLOv8 的交通标志检测算法研究[J]. 电子测量技术, 2024, 47(15): 117-124.
- [21] 朱强军, 胡斌, 汪慧兰, 等. 基于轻量化 YOLOv8s 交通标志的检测[J]. 图学学报, 2024, 45(3): 422-432.
- [22] ZHAO G M, GE W F, YU Y Z. GraphFPN: graph feature pyramid network for object detection[C]// International Conference on Computer Vision. Piscataway: IEEE, 2021: 2743-2752.
- [23] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection [C]// Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 10778-10787.
- [24] LI H L, LI J, WEI H B, et al. Slim-neck by GSConv: a lightweight-design for real-time detector architectures [EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2206.02424v3>.
- [25] YANG G Y, LEI J, ZHU Z K, et al. AFPN: asymptotic feature pyramid network for object detection[C]// 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Piscataway: IEEE, 2023: 2184-2189.
- [26] ZHANG H, ZHANG S J. Shape-IoU: more accurate metric considering bounding box shape and scale[EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2312.17663>.
- [27] ZHANG H, XU C, ZHANG S J. Inner-IoU: more effective intersection over union loss with auxiliary bounding box[EB/OL]. [2024-09-01]. <https://arxiv.org/pdf/2311.02877v4>.

(责任编辑:任滢滢)